APPENDIX A

TECHNICAL NOTES

General Information

The data in this report come from many sources, including surveys conducted by Federal and State agencies and by professional associations. The data reflect many methods of collection, such as universe surveys, sample surveys, and compilations of administrative records. Users should take great care when comparing data from different sources. Data often will not be strictly comparable due to differences in definitions, survey procedures, phrasing of questions, and so forth.

Survey accuracy is determined by the joint effects of "sampling" and "nonsampling" errors. Sampling errors arise because estimates based on a sample will differ from the figures that would have been obtained if a complete census had been taken.

All surveys, whether universe or sample, are also subject to nonsampling errors, which can arise from design, reporting, and processing errors as well as errors due to faulty response or nonresponse. These nonsampling errors include respondent-based events such as some respondents interpreting questions differently from other respondents; respondents making estimates rather than giving actual data; and respondents unable or unwilling to provide complete, correct information. Errors can also arise during the processing of responses, such as faulty imputation or reweighting to adjust for nonresponse, and recording and keying errors.

Racial/Ethnic Information

Data collection and reporting of the race/ethnicity of individuals pose several additional problems. First, both the naming of population subgroups and their definitions have often changed over time. Because this report draws on data from many sources, different terminology may have been used to obtain the various statistics presented here. Efforts have been made to maintain consistency throughout this text, but in some data reporting, it has been necessary to use distinct terminology that does not match other compilations.

Second, many of the groups of particular interest are quite small, so that it is difficult to measure them accurately without universe surveys. In some instances sample surveys may not have been of sufficient scope to permit calculation of reliable racial/ethnic population estimates, so that results are not shown for all groups. In addition, the reader is cautioned that it is easy to overlook or minimize the heterogeneity within subgroups when only a single statistic is reported for the total racial/ethnic group.

Information About Persons With Disabilities

The data on persons with disabilities in science and engineering are seriously limited for several reasons. First, the operational definitions of "disability" vary and include a wide range of physical and mental conditions. Different sets of data have used different definitions and thus are not totally comparable. The Americans with Disabilities Act of 1990 (ADA) encouraged progress toward standard definitions. Under the ADA, an individual is considered to have a disability if the person has a physical or mental impairment that substantially limits one or more of the major life activities, has a record of such impairment, or is regarded as having such an impairment. The ADA also contains definitions of specific disabilities. (See appendix table 1-1.)

Second, data about disabilities frequently are not included in comprehensive institutional records (e.g., in registrars' records in institutions of higher education). If included at all in institutional records, such information is likely to be kept only in confidential files at an office responsible for providing special services to students. Institutions are unlikely to have information regarding any persons with disabilities who have not requested special services. In the case of elementary/secondary school programs receiving funds to provide special education, however, counts for the entire student population identified as having special needs are centrally available.

The third limitation on information on persons with disabilities gathered from surveys is that it often is obtained from self-reported responses. Typically, respondents are asked if they have a disability and to specify what kind of disability it is. Resulting data, therefore, reflect individual perceptions, not objective measures.

Finally, data on persons with disabilities are often derived from sample surveys whose main purpose is to derive estimates for a full population. Deriving esti-

mates for any phenomenon that is applicable to a small proportion of the total is particularly difficult, especially when the sampling procedures do not have a way to "oversample" cases providing the characteristic of interest. Because persons with disabilities constitute a relatively small portion of the population, sample sizes may not be sufficiently large to permit calculation of reliable estimates.

An example in which these factors come together can be seen in the attempt to provide estimates of the proportion of the undergraduate student population with disabilities. Self-reported data from the undergraduate student population, queried on a survey to ascertain patterns of student financial aid, suggest that about 10 percent of the undergraduate population report having some disability; estimates from population surveys of higher education institutions, in contrast, place the estimate much lower, between 1 and 5 percent. Whether this discrepancy is the result of self-perception, incomplete reporting, nonevident disabilities, or differing definitions is difficult to ascertain.

Therefore, although considerable information is available on persons with disabilities and their status in the educational system and in the science and engineering workforce, it is often not possible to compare the numbers of persons with disabilities from different sources.

Primary Sources

Current Population Reports, P70-33: *Americans With Disabilities: 1991–92*

Contact:
Current Population Reports
Bureau of the Census
U.S. Department of Commerce
Washington, DC 20233
Tel: (301) 763-8300

This report presents data on the disability status of the noninstitutionalized population of the United States. The source of the data is a combined sample from the 1990 and 1991 panels of the Survey of Income and Program Participation. A supplement containing an extensive set of questions about disability status was included as part of the sixth wave of the 1990 panel and the third wave of the 1991 panel. Both of these waves were fielded between October 1991 and January 1992. The total sample size for this study was approximately 30,000 interviewed households. Estimation procedures were used to inflate weighted sample results to independent estimates of the civilian noninstitutional population of the United States.

Twelve questions were used to determine disability status for this study. These concerned the presence of limiting conditions such as difficulty with sensory and physical functional activities; difficulty with activities of daily living; the existence of specific conditions such as dyslexia, developmental disabilities, or other mental or emotional conditions; and the presence of a physical, mental, or other health condition limiting the kind or amount of work or housework that the person can do. For children, additional questions asked, for example, whether the children had received therapy or diagnostic services, had limitations in their ability to do regular schoolwork, or had a long-lasting condition that limited their ability to undertake activities such as walking and running. A person was considered to have a disability if the individual was identified affirmatively by any of the 12 category questions.

National Assessment of Educational Progress, 1969 to 1992

Contact:
National Center for Education Statistics
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208-5653
Tel: (202) 219-1761
Fax: (202) 219-1751

The National Assessment of Educational Progress (NAEP) is sponsored by the National Center for Education Statistics (NCES) and has been conducted since 1983 by the Educational Testing Service. The overall goal of the project is to determine the Nation's progress in education. Accordingly, NAEP encompasses a series of national sample surveys designed to assess students in 10 subject areas such as reading, mathematics, science, writing, and history. Begun in 1969, NAEP was conducted annually through 1980; since 1980 the project has been conducted biennially. NAEP has surveyed the educational accomplishments of 9-, 13-, and 17-year-old students (and, in recent years, those in grades 4, 8, and 12 as well). Over the years, NAEP has undergone extensive changes both in survey methodology and in the assessment areas covered, to reflect changing informational needs and possible changes in education achievement.

Since 1986, NAEP has included both main and long-term trend assessments. Both assessments use a complex multistage stratified sample of schools, selected to ensure adequate representation of schools with high enrollment of blacks and Hispanics. Both excluded students with limited English proficiency and students receiving special education services who were mainstreamed less than 50 percent of the time.

The 1992 main assessment estimated student achievement at a cross-sectional point in time. The cross-sectional samples used innovations in assessment methodology and populations definition. Approximately 1,200 schools and 26,700 students participated. Student

response rates ranged from 81 percent of students in grade 12 to 93 percent of students in grade 4.

The 1992 long-term trend assessment estimated the current status of achievement using the same sampling and assessment methodology used in previous years. Approximately 17,600 students in the combined age/grade level were tested in mathematics and in science. School response rates for the grade levels examined in 1992 ranged from 82 to 88 percent. Student response rates ranged from 83 to 94 percent.

American College Testing Program

Contact:
The American College Testing Program
2201 North Dodge Street
P.O. Box 168
Iowa City, IA 52243
Tel: (319) 337-1510

The American College Testing (ACT) Assessment is taken by college-bound high school students who request that the results be sent to designated colleges and scholarship boards. The ACT is designed to measure educational development in the areas of English, mathematics, social studies, and natural sciences. The test results are used in part to help predict how well students might perform in college. In 1994, approximately 892,000 students took the ACT examinations.

ACT standard scores are reported for each subject area on a scale from 1 to 36. A composite score is obtained by taking the simple average of the four standard scores and is an indication of a student's overall academic development across the four subject areas.

Since the 1984–1985 school year, national norms have been based on the most recent ACT test scores available from all students taking the test and who are scheduled to graduate in the spring of the year.

It should be noted that college-bound students who take the ACT Assessment are not, in some respects, representative of college-bound students nationally. First, students who live in the Midwest, South, and Rocky Mountains and Plains regions are overrepresented among ACT-tested students compared with college-bound students nationally. Second, ACT-tested students tend to enroll in public colleges and universities more frequently than do college-bound students nationally.

Scholastic Aptitude Test (SAT)

Contact:
College Entrance Examination Board
Educational Testing Service
Princeton, NJ 08541
Tel: (609) 771-7600

The Admissions Testing Program of the College Board comprises a number of college admissions tests, including the Scholastic Aptitude Test (SAT). The SAT is taken by students who need the results to apply to a particular college or university or scholarship board. High school students participate in the testing program as sophomores, juniors, or seniors—some more than once during these 3 years. If they have taken the tests more than once, only the most recent scores are tabulated.

The SAT reports subscores in the areas of mathematics and verbal ability. Students may also elect to take Achievement Tests in any of 21 subject areas; these exams are generally taken by students who are applying to the more competitive schools. In 1994, approximately 1.1 million students took the SAT examination, and more than 200,000 took at least one Achievement Test.

In 1987 the College Board initiated a review of the Admissions Testing Program and made significant changes in the SAT Program in 1993–94. Through the January 1994 test administration, SAT Program tests included the SAT, the Test of Standard Written English (TSWE), and the Achievement Tests. Beginning in March 1994, the SAT program was revised into two formats: the SAT I: Reasoning Test (the mathematical and verbal sections, with revisions beginning in March 1994) and SAT II: Subject Tests (formerly known as the Achievement Tests, with the revisions beginning in May 1994).

The SAT results are not representative of high school students or college-bound students nationally since the sample is self-selected. In addition, public colleges in a number of states require that students applying for admission submit ACT scores rather than SAT scores; thus, the proportion of students taking the SAT in some states is very low.

The 1994 National Norms Study of the Cooperative Institutional Research Program

Contact:
Higher Education Research Institute
Graduate School of Education
University of California
320 Moore Hall
Los Angeles, CA 90024-1521
Tel: (310) 825-1925
Fax: (310) 206-2228

This series, initiated in 1966, provides national normative data on the characteristics of students attending American colleges and universities as first-time, full-time, first-year students. The series is a project of the Cooperative Institutional Research Program (CIRP), a national longitudinal study of the American higher education system sponsored by the American Council on Education and the Graduate School of Education at the University of California, Los Angeles.

Since 1972, the CIRP freshman surveys have been conducted by the Higher Education Research Institute at the University of California, Los Angeles. The 1994 CIRP freshman norms are based on the responses of

237,777 students at 461 of the Nation's 2- and 4-year colleges and universities, statistically adjusted to reflect the responses of the 1.5 million first-time, full-time students entering college as freshmen in fall 1994.

The 1994 Student Information Form is a student self-report questionnaire composed of 39 multiple choice items. The questionnaire obtains data from students in eight areas: academic skills and preparation; demographic trends; high school activities and experiences; educational and career plans; majors and careers; attitudes; student values; and means of financing education.

The CIRP National Norms Study sample is derived from students attending institutions that volunteered to participate in the study. Therefore, it is not a random sample of the U.S. population of higher education institutions and students. As a result, survey findings may not present trends in the Nation as a whole.

The Integrated Postsecondary Education Data System Survey: Fall Enrollment, Completions and Institutional Characteristics

Contact:
National Center for Education Statistics
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208-5652
Tel: (202) 219-1373
Fax: (202) 219-1679

The Integrated Postsecondary Education Data System (IPEDS) began in 1986 as a supplement to and replacement for the Higher Education General Information Survey (HEGIS), which began in 1966. HEGIS was an annual survey of institutions listed in the current NCES Education Directory of Colleges and Universities; IPEDS surveys all postsecondary institutions, including universities and colleges and the institutions that offer technical and vocational education. The higher education portion is a census of accredited 2- and 4-year colleges, whereas technical and vocational schools are surveyed on a sample basis.

IPEDS consists of several integrated components that obtain information on types of institutions where postsecondary education is available, student participants, programs offered and completed, and the human and financial resources involved in the delivery of postsecondary education. The components of IPEDS include surveys of institutional characteristics; fall enrollment of students, including their age and residence; fall enrollment in occupationally specific programs; completions; finance; staff; salaries of full-time instructional faculty; and academic libraries.

The IPEDS Institutional Characteristics survey provides the basis for the universe of institutions reported in the Education Directory of Colleges and Universities. The universe includes institutions that met certain accreditation criteria and offered at least a 1-year program of college-level studies leading toward a degree. Each fall, institutions listed in the previous year's directory are asked to update information on the characteristics of their schools.

The IPEDS Completions Survey replaces and extends the HEGIS Degrees and Other Formal Awards Conferred Survey. The Completions Survey is administered to a census of institutions offering degrees at the bachelor's level and above, all 2-year institutions, and a sample of less-than-2-year institutions.

The IPEDS Fall Enrollment Survey replaces and extends the previous HEGIS surveys of institutions of higher education.

Imputations were developed for institutions that provided incomplete racial/ethnic data. Some of these institutions had reported total degrees awarded but not racial/ethnic data. In these cases, NCES imputed data on the basis of an earlier response for each institution, if available. The percentage of imputed data for racial/ethnic categories in 1993 ranged from 0.6 percent to 1.7 percent for bachelor's degrees, and from 1.8 percent to 7.0 percent for master's degrees. Other institutions reported totals that were larger or smaller than the sum of the racial/ethnic components, or reported racial/ethnic data as unknown. In these cases, NCES distributed the difference among the racial/ethnic groups for that institution.

Survey of Earned Doctorates

Contact:
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1774
Fax: (703) 306-0510

The Survey of Earned Doctorates (SED) has been conducted annually since 1957, under contract by the National Research Council of the National Academy of Sciences, for the National Science Foundation, the U.S. Department of Education, the National Endowment for the Humanities, the National Institutes of Health, and the U.S. Department of Agriculture. This is a census survey of all recipients of research doctoral degrees such as PhD or D.Sc.; it excludes the recipients of first-professional degrees such as J.D. or M.D. Therefore, SED data are restricted to research doctorates.

Data for the SED are collected directly from individual doctorate recipients. The recipients are asked to provide information on the field and specialty of their degree, as well as their personal educational history, selected demographic data, and information on their postgraduate work and study plans. Approximately 95 percent of the annual cohort of doctorate recipients respond to the questionnaire, which is distributed

through the cooperation of the graduate deans at institutions awarding doctorates.

Partial data from public sources, such as field of study, are added to the file for nonrespondents. No imputations are made, however, for nonresponse for data not available elsewhere, such as race/ethnicity information. The data for a given year include all doctorates awarded in the 12-month period ending on June 30 of that year.

Survey of Graduate Students and Postdoctorates in Science and Engineering: 1993

Contact:
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1774
Fax: (703) 306-0510

This annual survey collects data from all institutions offering graduate programs in any science, engineering, or health field. Data are collected at the academic department level. Available information includes full-time graduate students by source and mechanism of support, including data on women and first-year students enrolled full time; part-time graduate students by sex; and citizenship and racial/ethnic background of all graduate students. In addition, detailed data on postdoctorates are available by source of support, sex, and citizenship, including separate data on those holding first-professional doctorates in the health fields; summary information on other doctorate nonfaculty research personnel is also included.

In fall 1993, the latest survey cycle for which final data are available, the survey universe included approximately 11,150 departments at 605 institutions of higher education, including 346 doctorate- and 259 master's-granting institutions. Separate data were obtained from 120 specialized entities such as medical and dental schools, schools of public health, and other organizational units, bringing the total number of responding entities to 725. Coverage included all departments in 62 science, engineering, and health fields: 39 science fields (4 physical, 4 environmental, 1 mathematical, 1 computer, 1 agricultural, 17 biological, 1 psychology, and 10 social), 14 engineering fields, and 9 health fields.

The National Postsecondary Student Aid Study, "Undergraduate Financing of Postsecondary Education," 1992–93

Contact:
National Center for Education Statistics
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208-5652
Tel: (202) 219-1839
Fax: (202) 219-1736

The National Postsecondary Student Aid Study (NPSAS) was established by NCES to collect information concerning financial aid allocated to students enrolled in U.S. postsecondary institutions. After a national field test in 1985–1986, NPSAS was first administered in the fall of the 1986–1987 academic year. NCES conducted a second cycle of NPSAS for the 1989–1990 school year. This second cycle also contained enhancements to the methodology used in the 1987 cycle. The 1993 estimates, although generally comparable to the 1990 estimates, are not comparable to published estimates from the 1987 NPSAS.

The 1992–1993 in-school sample involved about 78,000 undergraduate and graduate students selected from registrar lists of enrollees at postsecondary institutions. The sample included students who did and did not receive financial aid. Student information such as field of study, educational level, and attendance status (part-time or full-time) was obtained from registrar records. Types and amounts of financial aid and family financial characteristics were abstracted from school financial aid records. Parents of students were also sampled to compile data concerning family composition and parental financial characteristics. Biennial follow-up data collections are expected. Students enrolled in postsecondary education for the first time in 1990 will serve as the base for the longitudinal component of NPSAS.

The SESTAT Data System

Contact:
Science and Engineering Personnel Program (PER)
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1776
Fax: (703) 306-0510

In the 1990s, NSF redesigned its data system about scientists and engineers. Termed SESTAT, the new data system integrates data from the NSF demographic surveys (Survey of Doctorate Recipients, National Survey of College Graduates, National Survey of Recent College Graduates), the Occupational Employment Statistics (OES) Survey, and administrative data from the Immigration and Naturalization Service (INS). The integration of the SESTAT demographic surveys requires complementary sample populations and reference periods, matching survey questions and procedures, as well as weighting adjustments for any overlapping populations.

The demographic surveys provide data on educational background, occupation, employment, and demographic characteristics. These surveys are of individuals and have a combined sample size of about 140,000. The OES, a large-scale survey of establishments, has occupational estimates by detailed industry category. The INS information provides counts of persons who have received permanent visas and who listed science or engineering as their occupation. OES and INS counts also include estimates of science and engineering technicians and technologists.

Scholars and policy analysts may access the SESTAT system through a variety of means, including access through the World Wide Web and restricted use data files. Individuals interested in obtaining more information about accessing the system should contact the Division of Science Resources Studies' Science and Engineering Personnel Program (PER) listed above.

Survey of Doctorate Recipients: 1993

Contact:
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1776
Fax: (703) 306-0510

The Survey of Doctorate Recipients (SDR) is a longitudinal survey designed to provide demographic and career history information about individuals with doctoral degrees. The survey is conducted for the National Science Foundation and other Federal agencies under contract by the National Research Council of the National Academy of Sciences. The 1993 survey, the 11th in a biennial series, reflects a number of improvements made by the National Science Foundation. The SDR is a survey of individuals under the age of 76 who hold doctorates in science and engineering from U.S. institutions. Several improvements introduced into the 1993 SDR affect comparability with SDR data published in prior survey years.

Among the variables included in this survey are citizenship, date of birth, disability status, educational history, employment status (unemployed, employed part time, or employed full time), field of degrees, geographic place of employment, labor force status, occupation, postdoctorate status, primary work activity (e.g., teaching, basic research), race/ethnicity, salary, sector of employment (academia, industry, government), sex, and years of professional experience.

The sample size for the 1993 survey was approximately 50,000 and had a response rate of 87 percent. The sample was stratified on the basis of field of degree, sex, disability status, racial/ethnic group, and nativity (i.e., whether born in the United States) to provide more reliable data on rare subgroups in the population. The sample frame used to identify these individuals is the Doctorate Records File, maintained by the National Academy of Sciences. The primary source of information for the frame is the Survey of Earned Doctorates (SED) (discussed separately above). For individuals who received a degree prior to 1957 when the SED started, information was taken from a register of highly qualified scientists and engineers that the National Academy of Sciences assembled from a variety of sources.

Because this is a longitudinal survey, recent recipients of research doctorates are added each time the survey is conducted and individuals no longer under age 76 are dropped. Initial data collection in 1993 was by mail. Nonrespondents to the mail questionnaire were followed up, using computer-assisted telephone interviewing techniques. The instrument used in the phone follow-up was modified from the mail instrument to avoid difficulties encountered in administering some of the questions by phone, especially those (such as field of degree and field of occupation) that require individuals to select from an extensive list of possible responses.

National Survey of College Graduates, 1993

Contact:
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1776
Fax: (703) 306-0510

The 1993 National Survey of College Graduates (NSCG) is a National Science Foundation survey of 215,000 individuals under age 76 who had a bachelor's degree or higher at the time of the 1990 Decennial Census. It is the primary source of data at NSF on scientists and engineers with bachelor's and master's degrees. The NSCG collects information on fields and levels of education, occupation, work activities, earnings, demographic, and other information on the science and engineering workforce.

The NSCG also contains information on PhDs, albeit with a much smaller sample size than NSF's biennial Survey of Doctoral Recipients (SDR), which remains the primary source of data on PhD scientists and engineers. The NSCG mailed to 10,000 individuals who had reported PhDs on the 1990 census and also picked up information on many individuals reporting bachelor's or master's degrees in 1990 who had completed the PhD by April 1993. An advantage of a sample drawn from the U.S. Census Bureau is that the NSCG includes data on PhDs and other degrees received from foreign institutions. Microdata on this part of the science and engineering population is not available from any other source. This survey is designed to be a baseline survey for the decade of the 1990s. Current plans are to follow individuals identified in this survey as having a science and engineering degree and/or a science or engineering occupation biennially between 1995 and 2001.

The definition of the population surveyed has changed several times over time. For example, the baseline survey in 1982 selected individuals with 4 or more years of college and did not screen for age. The changes between 1993 and prior surveys in population definition and other aspects of the survey are sufficiently great that NSF does not believe that meaningful trend analyses can be performed, comparing the 1993 data with the 1980s data.

National Survey of Recent College Graduates, 1993

Contact:
Division of Science Resources Studies
National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
Tel: (703) 306-1776
Fax: (703) 306-0510

The National Survey of Recent College Graduates (NSRCG) provides information about individuals who recently obtained bachelor's or master's degrees in a science or engineering field. Key variables include demographic information, employment status, field of degree, school enrollment status, occupation, sector of employment, primary work activity, salary, and years of professional experience.

The population of the 1993 survey consisted of all individuals under the age of 76 who received bachelor's or master's degrees in science or engineering between April 1, 1990, and June 30, 1992, from a U.S. institution. This survey is designed in part to cover individuals excluded from the National Survey of College Graduates, because they did not have a college degree as of April 1, 1990. Current plans are to follow a sample of individuals identified in this survey biennially between 1995 and 2001, along with individuals identified as being of interest in the NSCG.

The NSRCG sample is a two-stage sample. The first stage consists of selecting U.S. institutions that grant bachelor's or master's degrees in science and/or engineering fields. The sample frame of schools for inclusion in the first stage of the sample is obtained from the Integrated Postsecondary Education Data System database maintained by the National Center for Education Statistics. In 1993, 274 institutions were selected in the first-stage sample. The sample frame for the selection of graduates is obtained from representatives of the institutions selected at the first stage. In total, 26,000 individuals were selected in 1993. The current estimated response rate for the first stage of this survey in 1993 is approximately 99 percent and for the second stage is approximately 86 percent.

A number of changes have been made in the definition of the population surveyed over time. For example, the 1990 survey included individuals receiving bachelor's degrees in fields such as engineering technology; these are excluded from the 1993 survey. The changes between 1993 and prior surveys in population definition and other aspects of the survey are sufficiently great that Science Resources Studies staff believe that trend analyses must be performed very cautiously, if at all.